

Resumen

Este escrito pretende relacionar algunas razones por las cuales se considera importante tener presente la ética como piedra angular en la implementación de la Inteligencia Artificial (IA) y sus productos, para ello, se relacionarán las posturas de algunos autores expertos y organismos internacionales respecto de los riesgos de la automatización de esta y los potenciales efectos de sus productos. De igual forma, se abordará el transhumanismo como un mecanismo que propende por la optimización de la calidad de vida de los seres humanos, pero siempre bajo el acatamiento de los derechos humanos y su dignidad.

Palabras clave

Inteligencia Artificial, ética, heurística, algoritmo, autodeterminación, transhumanismo.

Abstract

This writing intends to leave evidence of the importance of keeping ethics in mind as a cornerstone in the development of Artificial Intelligence IA, for this, the positions of some expert authors and international organizations regarding the risks of automation of this and the potential effects of their products. Similarly, transhumanism will be addressed as a mechanism that tends to improve the quality of life of humans under the respect of human rights and their dignity in the use of AI.

Keywords

Artificial Intelligence, ethics, heuristics, self-determination, algorithm, transhumanism.

Recepción: 13.09.2022

Aceptación: 16.01.2023

Cite este artículo como:

Zabala, T. D. (2023). La ética en la inteligencia artificial. (M. M. Quiroz, & D. Zamora, Edits.) *Revista Gestión de la Seguridad y Salud en el Trabajo*, 5(2 (7)), 82 - 86.

¹ Resultado del proyecto de investigación “Desafíos de la regulación de la inteligencia artificial y sus aplicaciones desde la perspectiva del Derecho”; Universidad Politécnico Grancolombiano.

² Profesional en Derecho; Magíster en Derecho; tzabala@poligran.edu.co; Filiación institucional Politécnico Grancolombiano; Programa Técnico Profesional Judicial Modalidad Virtual.

Introducción

El veloz desarrollo de la Inteligencia Artificial ha generado temor en quienes la estudian y la aplican, puesto que ella ha venido adquiriendo nuevos niveles de automatización que le están permitiendo incluso llegar a tomar acciones complejas sin la mediación del humano. Un evento que podría mencionarse es el de los coches autónomos, que tienen la capacidad de determinarse frente a los riesgos propios de la conducción a través de procesos heurísticos para la resolución de conflictos (desde los más básicos a los más complejos), lo que implica dilemas éticos de gran significancia como la decisión de aplicar procesos óptimos o lógicos en eventos de riesgo a la vida e integridad de las personas (Marín S., 2019).

De acuerdo con (Bonnefon, 2016), la IA debe contar con algoritmos con patrones de conducta ética que propendan por la garantía de la integridad de los individuos, no obstante, existe el riesgo de conflicto cuando en un mismo evento están en riesgo la vida o la integridad del usuario de la IA y un tercero, pues en dicho caso ¿Cuál derecho prima más? O ¿Quién está legitimado para tener un derecho preferente cuando son sujetos protegidos en el mismo nivel?, etc.

Existe al respecto una propuesta bastante interesante, (Martín. K. y Freeman. R., 2004) consideran que es un error común considerar que los avances tecnológicos siempre son buenos, puesto que toda innovación o evolución artificial tiene sus riesgos para el entorno en el que han de desempeñarse. Para hacer una reflexión sobre el alcance de la IA y sus productos, vale la pena mencionar que estas están divididas en 4 categorías de definición según sus objetivos como lo son: 1) aquellas que simulan el comportamiento humano, 2) las que tienen sistemas cognitivos similares al de los humanos, 3) las que simulan los procesos de razonamiento y 4) las que realizan actos racionalmente (Russell S. y., 2016).

En ese sentido, es posible aseverar que la principal pretensión de la IA es representar la inteligencia humana, pero si a ello se le suman los objetivos de su funcionamiento también podría considerarse que estas buscan actuar como los humanos y con ello producir efectos jurídicos tan complejos como las que surgen de las interacciones de los individuos en sociedad.

Marco Teórico

Partiendo del hecho de que la ética es aplicable a los sujetos dotados de autonomía y que, según la legislación vigente, esta es exclusiva del ser humano por estar dotado de inteligencia, autodeterminación, raciocinio y libertad en su concepción más amplia ((EGE), 2018), sería objetable hacer extensiva la ética a la IA, por no ser esta un sujeto autónomo y racional. No obstante, (Quiroz, 2014) sostiene que la IA tiene la capacidad de funcionar de manera independiente, pero sin poder dar cuenta de sus propios actos o tomar disposiciones con la totalidad de los elementos inherentes al humano y provenientes de su propia conciencia. Al respecto (Marín Palma, 2019), considera que es erróneo concebirlos como autónomos debido a que no deben actuar sin supervisión humana, por lo que el término más adecuado es el de automáticos y no autónomos. Siendo así, la automaticidad está referida a la independencia que adquieren la IA en su operatividad al poder actuar dentro de los parámetros que le han sido autorizados en su programación sin tener que requerir permisos para su cotidiano funcionamiento (con algunas excepciones).

Ahora bien, (López de Mantaras, 2015) clasifica la IA en dos modalidades, una fuerte y otra débil. Por su parte, la débil es la que se usa para el funcionamiento de los aparatos que requieren la intervención permanente del humano como los computadores, los celulares, Siri, los asistentes de Amazon, etc., y los fuertes que son aquellos que simulan el pensamiento y el comportamiento humano, siendo estos últimos los más cercanos a los autónomos. Teniendo presente que el nivel de automatización de la IA cada vez es más evidente y permisivo, si es recomendable que los diseñadores apliquen principios éticos en la concepción, desarrollo y aplicación de la IA, pues esta de ninguna manera debe perjudicar a los seres vivos que con ella interactúan; y es justamente aquí, donde la concepción del transhumanismo adquiere un papel fundamental; entendiéndose este como aquel que pone en cuestionamiento la transformación y readaptación de la condición humana a través del uso de la tecnología para detener el envejecimiento, mejorar las capacidades biológicas y naturales, aumentar la capacidad mental o cognitiva, etc., es decir, para mejorar la raza humana. Siendo así, al implantar en el cuerpo orgánico humano tecnología

o IA podría darse rienda suelta a la autonomía con todo lo que ello implica, pues el determinante de los actos sería el individuo, quien es capaz de legislarse a sí mismos (Russell J. y., 2012).

Es así como la Unión Europea, través de su Parlamento y su Comisión, siendo la organización internacional más avanzada en el análisis y regulación de la IA y sus productos, ha propuesto que debe implementarse la ética partiendo de la dignidad humana (justicia, bien común, prosperidad, no discriminación, paz y equidad), derechos humanos y los derechos fundamentales (Carta Internacional de los Derechos Humanos) (Parlamento Europeo, 2017) ((EGE), 2018).

Al respecto, un grupo de expertos en Inteligencia Artificial conocido como AI-HLEG, (Floridi, 2018) propuso como objetivos del desarrollo de las tecnologías relacionadas con la IA y sus productos las siguientes: la protección de la democracia y el Estado de Derecho, mayor oferta de bienes y servicios con costos más bajos y calidad más alta, riesgos mínimos y mitigación de daños (parámetros robustos para la prevención de perjuicios), confianza del usuario y medio ambiente sostenible. En ese orden de ideas, según estos objetivos, estaría prohibido el uso armamentístico de la IA y, el ser humano, en ninguna circunstancia podrá perder el poder de decisión sobre la IA y sus productos, además de reservarse la facultad de supervisión permanente sobre esta con la posibilidad de cambiar los patrones que evidencien potencial riesgo de daño, es decir, control humano sin importar el nivel de desarrollo de la IA y sus productos (Floridi, 2018).

Es por ello por lo que el Parlamento Europeo insiste en la necesidad de crear un Código de ética que incluya temas como: definición de jurisdicción y competencia en el asunto, prevención del conflicto de intereses, perfil de experiencia e idoneidad de los desarrolladores de la IA, parámetros para la rendición de cuentas, transparencia, multidisciplinariedad y la obligación de los Estados de educar a su población sobre la IA, sus alcances y sus usos.

El autor (Asís, 2020) afirma que no existen axiomas éticos en la evolución de las ciencias y tecnologías que limiten los problemas sociales, haciendo que los derechos humanos destaquen en términos de

amparo de las garantías como los valores universales.

Por lo tanto, la Declaración de Toronto publicada en mayo de 2018 enfatiza tres aspectos para aprovechar el uso de la regulación y producción de IA: el primero está relacionado con la no discriminación y la prevención de la segregación en el diseño e implementación de sistemas de aprendizaje automatizado para entornos públicos; el segundo corresponde a procesos de rendición de cuentas para desarrolladores, implementadores y usuarios de IA; y el tercero es la implementación de auditorías eficientes y eficaces para el seguimiento, reconocimiento e individualización de los infractores de los derechos de las personas a través de la IA y/o sus productos (Asís, 2020).

En este sentido, la Declaración de Toronto busca proteger los derechos individuales impulsando modelos e IA con educación inclusiva para aplicar sistemas de aprendizaje automatizados que brinde los espacios para que las personas con necesidades especiales o vulnerabilidades accedan a los servicios y garantías disponibles para los demás con justicia e igualdad (Naciones Unidas, 2018). Por su parte, (*The public voice*, 2018) organismo internacional, realizó un aporte significativo al sugerir los Lineamientos Universales para el desarrollo, ejecución y uso de la IA, encaminados a racionalizar los beneficios que ofrece a las personas naturales bajo criterios de minimización de riesgos y protección de los Derechos Humanos, así:

- ✓ Justicia e igualdad
- ✓ Transparencia
- ✓ Comprensión - capacidad para explicarse e interpretarse
- ✓ Responsabilidad en las acciones
- ✓ No maleficencia
- ✓ Beneficencia
- ✓ Privacidad
- ✓ Sustentabilidad
- ✓ Confianza
- ✓ Dignidad
- ✓ Solidaridad
- ✓ Cohesión y seguridad social

Marco Metodológico

Este artículo es producto de un estudio de tipo cualitativo por ser el medio más idóneo para la recolección de la percepción de expertos, cuyo



resultado se dio en una investigación exploratoria (Hernández, 2010).

Resultados

Desde una perspectiva jurídica en Colombia, hasta la fecha, el tema de la IA, sus productos y su código de ética no han sido objeto de estudio o por lo menos, considerados como objeto de regulación por el legislativo, siendo así, todo lo relacionado con la responsabilidad civil contractual y extracontractual que de ella derivada se le asigna a quien sea su desarrollador o propietario, obviando el alcance de la automatización de la IA.

Adicional a lo antes dicho, para iniciar la proyección de regulación es menester definir si la IA y sus productos son o no sujetos de derechos y obligaciones y en qué nivel se encuentran respecto del humano, cuál es el alcance de su responsabilidad, la determinación de sus atributos y si su operatividad clasifica o no dentro de los actos jurídicos; pues sería este el medio idóneo para asignárseles normatividad exclusiva en materia de responsabilidad civil, penal, disciplinaria, entre otras.

En lo que al transhumanismo se refiere, este debería circunscribirse a los parámetros legales aplicables al humano, pues sería el individuo quien determinaría el uso de los elementos tecnológicos que le han sido implantados y no al contrario, lo que quiere decir, que debería prohibirse la implantación de elementos tecnológicos que limiten o nublen la autodeterminación de la persona.

Conclusiones

Es por ello, que sí se hace necesario regular el diseño, desarrollo e implementación de la IA y sus productos, establecer un código de ética para la prevención de la mala *praxis* y la extralimitación del uso de las tecnologías, prohibir expresamente el uso armamentístico, el uso de datos personales, la aplicación de cláusulas abusivas y limitar el alcance de la perspectiva dominante en las interacciones con los seres vivos.

Así mismo, la libertad y la autonomía debe mantenerse como característica exclusiva del ser humano y someterla a control permanente. Su operatividad debe estar parametrizada dentro del respeto por la vida, la dignidad, la justicia, la

equidad, la no discriminación, la paz y la prohibición del monopolio en la *praxis* de las ciencias, disciplinas, ocupaciones u oficios.

Referencias

- [1] Asís, R. (2020). *Inteligencia artificial y derechos humanos: Seminario Permanente Gregorio Peces-Barba, Materiales de Filosofía del Derecho*. 20-04. Universidad Carlos III de Madrid. <https://e-archivo.uc3m.es/handle/10016/30453>.
- [2] Bonnefon, J. S. (2016). The social dilemma of autonomous vehicles. *Journal Science*, 352(6293), 1573-1576. doi: 10.1126/science.aaf2654.
- [3] European Group on Ethics in Science and New Technologies (EGE). (2018). *Statement on Artificial Intelligence, Robotics and “Autonomous” Systems*. Comisión Europea, Dirección General de Investigación e Innovación. <https://op.europa.eu/es/publication-detail/-/publication/dfebe62e-4ce9-11e8-be1d-01aa75ed71a1/language-en/format-PDF/source-78120382>.
- [4] Floridi, L. E. (2018). AI4People —An Ethic. Grupo de Expertos de Alto Nivel en Inteligencia Artificial (AI-HLEG) (2018). Draft ethics guidelines for trustworthy AI: Working document for stakeholders’ consultation. <https://www.incp.org.co/Site/publicaciones/info/archivos/AIHLEG-Draft-Ethics-Guidelines.pdf>.
- [5] Hernández, R. F. (2010). *Metodología de la investigación*: Mc Graw Hill. <https://www.icmujeres.gob.mx/wp-content/uploads/2020/05/Sampieri.Met.Inv.pdf>.
- [6] López de Mantaras, R. (2015). Algunas reflexiones sobre el presente y futuro de la inteligencia artificial. *Novática*, 234(4), 97-101. <https://digital.csic.es/handle/10261/136978>.
- [7] Marín, R. (2019). *Inteligencia artificial. Técnicas, métodos y aplicaciones*. McGRAW-Hill/Interamericana De España, S.A.U. <https://editorial.tirant.com/es/libro/inteligencia-artificial-tecnicas-metodos-y-aplicaciones-9788448156183>.
- [8] Marín, S. (2019). Ética e Inteligencia Artificial. *Cuadernos de la Cátedra CaixaBank de Responsabilidad Social Corporativa*. (42). <https://www.expoelearning.com/wp-content/uploads/2020/01/%C3%89tica-e-inteligencia-artificial.pdf>.

- [9] Martín. K. y Freeman. R. (2004). The Separation of Technology and Ethics in Business Ethics. *Journal of Business Ethics*, 53(4),353-364. <https://link.springer.com/article/10.1023/B:BUSI.000043492.42150.b6>.
- [10] Naciones Unidas. (2018). *The Toronto Declaration: Protecting the right to equality and non-discrimination in machine learning systems*. https://www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration_ENG_08-2018.pdf.
- [11] Parlamento Europeo. (2017). *Resolución del Parlamento Europeo. Recomendaciones destinadas a la Comisión sobre normas de derecho civil sobre robótica* (2015/2103[INL]). https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_ES.html.
- [12] Quiroz, M. (2014). Acercamiento a las “oposiciones paradigmáticas” entre neoconstitucionalismo y positivismo jurídico. *Revista IUSTA*, 41(2), 77-97. <https://www.redalyc.org/pdf/5603/560358701004.pdf>.
- [13] Rusell, J. Y. (2012). *YAML*. Bookvika Publishing.
- [14] Rusell, S. y. (2016). *Artificial Intelligence: A Modern Approach* (3.ª edición). Malasia: Pearson Education Limited. <https://web.cs.ucla.edu/~srinath/static/pdfs/AIMA.pdf>.
- [15] The public voice. (23 de octubre de 2018). *Directrices universales para la inteligencia artificial*. <https://thepublicvoice.org/ai-universal-guidelines/>.